Casey Sheahan

LDES 504

Evidence of Learning Final Project

5/19/2019

**Explanation of Data**

For this project, I decided to analyze some of our admissions data to get a better understanding of our applicants. The School of Medicine receives around 14,000 applications a year, but interviews under a thousand. For the sake of this project and to get a better understanding of the applicants that get through to the next step, I decided to look at only the applicants that we interview for the MD program. I hope to use this information for recruitment purposes, so we have a better idea of where the applicants we are interviewing are coming from, especially our top applicants. For the four graphs I created I used GPA, State, MCAT score, and Gender to find trends. While we do have international students, the large majority of our applicants are from the United States, so I decided State would still be helpful. We also collect information on Gender Identity, but it is an optional field and the large majority of our applicants don’t fill out this field. Since we do have “Decline to Answer” as an option for gender, I decided to use the Gender field instead of incorporating the Gender Identity data just for these specific graphs, but I took time to think through that decision. Now that I have explained a bit of how I went about thinking through the data I was going to analyze, I’m now going to move into my graphs and the specifics of some trends I can pull from them.

**Graph 1: State vs. GPA**

I first decided to see the relationship between the State of the applicant’s permanent address and their GPA. My first attempt at this graph was quite a wreck. I’ve included it below just because it was a funny starting place. When this is the first graph that R generated, I knew I had a way to go, but I eventually was able to correctly code the graph and created the second graph below.

This graph shows the GPAs of students from the States they applied from. I also used the color of the dots to designate if the dot represents a female, male, or an applicant that declined to answer. While this graph isn’t the most helpful, it does a great job of showing outliers and the general range of GPAs per state. It is also interesting to see that the mix of female and male seems pretty equal across the graph. I think this graph could have been improved by finding the average of the GPA of each state and to just graph those points. It would be much easier to see the different dots and how they compare across the graph that way, but I do think it’s really interesting to see the male and female aspect that this graph shows. I also think that the general break down of ranges and how many dots fall between 3.0 – 3.2 vs 3.6-4.0 is not unexpected, but interesting! Also when this graph is expanded in R it is much easier to read the states on the y-axis.

**Graph 2: GPA compared to MCAT**

 This was actually one of my graphs to study because the correlation between MCAT and GPA is one that is discussed very often in Medical School Admissions. Many applicants believe that GPA is not a great determinant of a student’s academic ability and other applicants also believe the exact opposite, which is that GPA is a good determinant, but MCAT score is not. I was really excited to see the correlation between the two, if any. While there are outliers from almost every GPA, you can see a small incline in the top MCAT scores as the GPA goes up. It seems that around 3.9-3.9 is the range where applicants have the highest MCAT scores. Another interesting piece of information that can be drawn from the graphs is we don’t see a really low MCAT score until around an applicant that has a 3.5 GPA. It does make sense because an applicant wouldn’t be invited for an interview with a low GPA, unless the rest of their applicant was outstanding, but to see where the actual cut off was where low MCATs are outweighed by the rest of their application is fascinating. Our office holistically reviews applications, so one lower aspect of their application can be outweighed by the rest of it. The other parts of the application that are weighed, but couldn’t be graphed in R are their essays, service and clinical experience.

**Graph 3: Amount of Applicants per State**

 I used the bar chart to graph the amount of applicants we interviewed from each state. It was also very helpful in this chart to use colors to distinguish between the gender of the applicants. I also want to mention when this graph is expanded in R Studio, the states are much more readable. I also turned the text by 90 degrees, so that they wouldn’t be overlapping on the x-axis. Interestingly enough, California has been the state where we receive most of our applicants from for the past couple years, so once again California was state were we had the most amount applications from and interviewed the most from. It also seems that most states have close to the same amount of males and females being interviewed. When we read applications, we actually don’t take into account what state the applicants are from since we are a private institution, so it’s interesting that the gender of the applicants from each state is pretty equal. This graph also helps show which states we should focus on a bit more when recruiting because of the low amount of applicants. It isn’t very surprised that the DMV area also ranks very high for amount of applicants.

**Graph 4: Average MCAT Score per State**

I used the boxplot chart to graph the average MCAT score per state and found this to be a perfect graph to chart this data for a couple of reasons. First of all it does a great job of showing the average MCAT per score and I also find it incredibly helpful that it denotes the outliers by showing them represented by a dot. Most of the average MCAT scores seem to be around the same average of about a 512-515. For the 2018 application cycle, our average MCAT was a 511, so it seems just from looking at this graph that the average may have gone up a point or two this year. It also might be worth it to put more effort into recruiting from some of the states that seem to have above average MCAT scores and maybe even look into the couple outlier applications to see how strong the rest of their application really was. The boxplot chart would also do a better job than the dot chart which I used to see GPAs per state, so that is a graph I want to create and would also be really helpful for recruiting.

Conclusion

Overall, I found this final project really interesting especially since I was able to use data from GUSOM Office of Admissions. There are so many other ideas I have for what to chart and different graphs that would be helpful. I also look forward to graphing data from 200 applicants that make up the final incoming class. Using graphs in R was my favorite thing to learn about, so it really felt full circle to be able to use it for something that can positively impact my office and full time job.